



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/647,709	08/25/2003	Senis Busayapongchai	60027.0303US01/BS030022	2462
23552 7590 05/21/2007 MERCHANT & GOULD PC P.O. BOX 2903 MINNEAPOLIS, MN 55402-0903			EXAMINER SHAH, PARAS D	
			ART UNIT 2609	PAPER NUMBER
			MAIL DATE 05/21/2007	DELIVERY MODE PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary

Application No.

10/647,709

Applicant(s)

BUSAYAPONGCHAI, SENIS

Examiner

Paras Shah

Art Unit

2609

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 04/20/2007.
- 2a) ☒ This action is **FINAL**. 2b) ☐ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-11, 13, 15-17 and 22-24 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1-11, 13, 15-17, and 22-24 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 20 April 2007 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
- Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
 - ☐ Certified copies of the priority documents have been received in Application No. _____.
 - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- ☐ Notice of References Cited (PTO-892)
- ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- ☐ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____
- ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____
- ☐ Notice of Informal Patent Application
- ☐ Other: _____

DETAILED ACTION

1. This Office Action is in response to the Amendment filed on 04/20/2007. Claims 1-11, 13, 15-17, and 22-24 remain pending with claims 12,14, and 18-21 being cancelled. The Applicants' amendment and remarks have been carefully considered, but they are not persuasive and do not place the claims in condition for allowance. Accordingly, this action has been made FINAL.
2. All previous objections and rejections directed to the Applicant's disclosure and claims not discussed in this Office Action have been withdrawn by the Examiner.

Response to Arguments

3. Applicant's arguments (pages 11-15) filed on 04/17/2007 with regard to claims 23 have been fully considered but they are not persuasive.

Applicant has argued that the limitation of recognizing the utterance by grammar type is not disclosed in the Kennewick *et al.* and the Crepy *et al.* reference. The Applicant has added this limitation from the dependent claim into the independent claim. The Examiner traverses the said arguments by showing a similar grammar categorizing is seen in the reference by Kennewick *et al.*, which is found in [0016], [0108], and [0144]. From these cited sections it is evident that the grammar is used to determine the key word. Once the keywords are identified, the context is determined as to what the query or command is related to. The example shown in the Kennewick *et al.* reference pertains to the recording of a TV program. It is implied that the words are identified and a query formulated depending on the context or the subject for which the command

Art Unit: 2609

inputted by the user. Further, the Knott *et al.* reference was cited for the grammar portion of the limitation, which identifies the command responses grouped together by use of a glossary (e.g. Also, interpreted to be a grammar for response types). The Kennewick reference was used for which the Applicants have considered and have denoted that there is no teaching of the limitation. From the below rejections and the previously said comments, it will be shown that Kennewick *et al.* and Knott *et al.* discloses a similar grammar sub-tree as well as the grammar sub-tree being grouped. The Applicant has also incorporated the limitation of loading the word or phrase into a memory location. The limitation is seen in the Crepy *et al.* reference in col. 2, lines 64-66 and the Kennewick *et al.* reference the following is implied that the word or phrases uttered is loaded into memory for context evaluation and speech recognition (see [0010] and [0012]).

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. The factual inquiries set forth in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:

1. Determining the scope and contents of the prior art.
2. Ascertaining the differences between the prior art and the claims at issue.
3. Resolving the level of ordinary skill in the pertinent art.
4. Considering objective evidence present in the application indicating obviousness or nonobviousness.

6. Claims 1-6, 8-11, 13, 15-17, and 22-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Kennewick *et al.* (US PGPub 2004/0044516) in view of Crepy *et al.* (US 6,622,121) in further view of Knott *et al.* (US PGPub 2003/0191648).

As to claim 1, Kennewick *et al.* discloses the improvement of speech recognition engine, comprising: identifying one or more utterances (see page 1, right column, [0010]; lines 3-4) for recognition by a speech recognition engine (see Figure 1, element 120 and page 6, right column, [0088], line 4); passing the one or more identified utterances to a text-to-speech module (see page 1, right column, [0012], lines 8-9 and Figure 1, element 124); a **grammar categorizing words and phrases from utterance** (see [0016], [0108], and [0144]) (e.g. From these cited sections it is evident that the

Art Unit: 2609

grammar is used to determine the key word. Once the keywords are identified, the context is determined as to what the query or command is related to. The example shown in the Kennewick *et al.* reference pertains to the recording of a TV program); analyzing each utterance from the **selected grammar sub tree** (see page 14, left column, [0188], line 8 and [0108]) (e.g. The context is analyzed to find out relevant subject matter) to determine how close each recognized utterance approximates to the audio pronunciation from each utterance derived (see page 14, left column, [0188], lines 8-10). However, Kennewick *et al.* does not specifically disclose the passing of the audio pronunciation of the identified utterance to the speech recognition engine and creating an utterance for each audio pronunciation that was passed and the loading of the word and phrases into memory, as well as the grouping of the in a grammar tree. Crepy *et al.* discloses the **loading of word and phrase into memory** (see col. 2, lines 64-66) and the passing of the audio pronunciation of each of the utterances to the speech recognition engine (see col. 1, line 65); creating an utterance for each audio pronunciation passed to the speech recognition engine (see col. 1, line 66-67). Knott *et al.* discloses the **grouping together of grammar type in a grammar sub-tree** (see page 3, right column, [0021], lines 5-7) (e.g. The following shows grouping of affirmations and refutation is similar to what the applicant is interpreting grammar type to be. Also, the groupings allow the tree for responses to be built, where negative answers and positive answers are known and selected upon user input). It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the identification of utterances and the analysis of the utterance by a

grammar and context presented by Kennewick *et al.* with the utilization of the output of the text to speech module into the speech recognition engine as presented by Crepy *et al.* The motivation to combine these two references involves testing the recognition of a spoken input (see Crepy *et al.*, Abstract). Further, the inclusion of the grouping of similar grammar type in a grammar tree allows various answers given by users (see Knott *et al.*, page 1, left column, [0003], lines 1-4), which would benefit the teachings of Kennewick *et al.* to include various utterances having similar meanings.

As to claim 2, Kennewick *et al.* discloses assigning a confidence score to each utterance (see Figure 5, element 506).

As to claim 3, Kennewick *et al.* discloses the assigning of confidence score to each recognized utterance based on a confidence level associated with the utterance based on prior speech recognition engine training (see page 10, left column, [0151], line 4-5).

As to claims 4 and 10, Kennewick *et al.* discloses the determination being made of whether the recognized utterance is the same as the utterance derived by the speech recognition engine based on prior speech recognition training **confidence level** (see page 10, right column, [0151], lines 4-8) (e.g. It should be noted that there is a dictionary that is used to see whether the recognized utterance matches). It is inherent that the words from the dictionary and the words from the utterance are matched for similarity.

As to claims 5 and 11, Kennewick *et al.* discloses wherein if the confidence score exceeds an acceptable level designating the recognized utterance as accurately

recognized by the speech recognition engine (see page 14, left column, [0188], lines 30-33).

As to claim 6, Kennewick *et al.* discloses wherein if the confidence score less than a certain value, a modification is made to the speech recognition engine to recognize the word (see page 14, left column, [0031]) (e.g. If the confidence level is less than a value, the system requests verification from a user or asks a question to remove any ambiguity. This is seen as a modification to the speech recognition engine to interpret the utterance).

As to claim 8, Kennewick *et al.* discloses whereby modifying the speech recognition engine includes altering the pronounced utterance associated with the confidence score that is less than a threshold value (see page 10, right column, [0153], lines 16-21) (e.g. The additional information obtained from the user removes the ambiguity in the uttered word due to low confidence. Thus, the uttered word is altered to be recognized) such that the altered audio pronunciation obtains an acceptable confidence score upon next pass (see page 10, right column, [0154], lines 17-19) (e.g. The confidence scores are updated as the system learns more information).

As to claim 9, Kennewick *et al.* discloses the reduction of the confidence score threshold level (see page 10, right column, [0154], lines 17-19). It is inherent that the constant update and learning of the system presented in the reference would alter the confidence score threshold as it would alter the confidence level of the word.

As to claim 13, Kennewick *et al.* discloses the extracting of one or more utterance via a dictionary unit (see page 10, left column, [0147], lines 2-4) (e.g. It should be noted that extraction is done by using the information from the dictionary).

As to claim 15, Knott *et al.* discloses the inclusion of the subcategories as a group containing all utterances (see page 3, right column, [0021], lines 5-7)) (e.g. The glossary contains both the refutations and affirmations).

As to claim 16, Knott *et al.* discloses the identifying of an utterance for recognition by identifying the category for which the spoken word belongs (see page 3, right column, [0021], lines 5-13) (e.g. It is inherent that depending when finding the confidence score the value of the correct category for which the spoken word is associated with is identified).

As to claim 17, Crepy *et al.* discloses the conversion of the audio pronunciation from audio format to a digital format (see col. 4, line 57-58) (e.g. The reference states that the conversion is done after text to speech synthesis. The conversion from audio to digital before the signal passes into the speech recognition engine or by the speech recognition engine (which is done before the recognition process) will have no effect on the result (utterance recognition)); and analyzing phonetically the audio pronunciation of the utterance to create the recognized word (see col. 4, line 59-64) (e.g. it should be noted that in order to compare the results from storage to that uttered, comparisons are done between the two. This would involve comparing the phonemes of the uttered word and the stored word).

As to claim 22, Kennewick *et al.* discloses the improvement of speech recognition engine, comprising: identifying one or more utterances (see page 1, right column, [0010], lines 3-4) for recognition by a speech recognition engine (see Figure, element 120 and page 6, right column, [0088], line 4); a **grammar categorizing words and phrases from utterance** (see [0016], [0108], and [0144]); passing the one or more identified utterances to a text-to-speech module (see page 1, right column, [0012], lines 8-9 and Figure 1, element 124) in a **selected grammar sub-tree** (see [0108]); analyzing each utterance (see page 14, left column, [0188], line 8) to determine how close each recognized utterance approximates to the audio pronunciation from each utterance derived (see page 14, left column, [0188], lines 8-10); assigning a confidence score to each recognized utterance based on speech recognition engine's confidence in each recognized utterance based on prior training of the speech recognition engine to recognize similar words (see page 10, right column, [0151], lines 4-8) (e.g. It should be noted that there is a dictionary that is used to see whether the recognized utterance matches); if the confidence score is less than an acceptable threshold, modifying the speech recognition engine to recognize the utterance (see page 14, left column, [0031]) (e.g. If the confidence level is less than a value, the system requests verification from a user or asks a question to remove any ambiguity. This is seen as a modification to the speech recognition engine to interpret the utterance). However, Kennewick *et al.* does not specifically disclose the deriving and passing of the audio pronunciation of the identified utterance to the speech recognition engine and creating an utterance for each audio pronunciation that was passed, as well as the grouping of the grammar types in a

Art Unit: 2609

grammar sub-tree. Crepy *et al.* discloses the passing of the audio pronunciation of each of the utterances to the speech recognition engine (see col. 1, line 65); creating an utterance for each audio pronunciation passed to the speech recognition engine (see col. 1, line 66-67). Knott *et al.* discloses the **grouping together of grammar type in a grammar sub-tree** (see page 3, right column, [0021], lines 5-7) (e.g. The following shows grouping of affirmations and refutation is similar to what the applicant is interpreting grammar type to be. Also, the groupings allow the tree for responses to be built, where negative answers and positive answers are known and selected upon user input). It would have been obvious to one of ordinary skilled in the art to have modified the identification of utterances and the analysis of the utterance presented by Kennewick *et al.* with the utilization of the output of the text to speech module into the speech recognition engine as presented by Crepy *et al.* The motivation to combine these two references involves testing the recognition of a spoken input (see Crepy *et al.*, Abstract). Further, the inclusion of the grouping of similar grammar type in a grammar tree allows various answers given by users (see Knott *et al.*, page 1, left column, [0003], lines 1-4), which would benefit the teachings of Kennewick *et al.* to include various utterances having similar meanings.

As to claim 23, Kennewick *et al.* discloses whereby modifying the speech recognition engine includes altering the pronounced utterance associated with the confidence score that is less than a threshold value (see page 10, right column, [0153], lines 16-21) (e.g. The additional information obtained from the user removes the ambiguity in the uttered word due to low confidence. Thus, the uttered word is altered to

Art Unit: 2609

be recognized) such that the altered audio pronunciation obtains an acceptable confidence score upon next pass (see page 10, right column, [0154], lines 17-19) (e.g. The confidence scores are updated as the system learns more information).

As to claim 24, Kennewick *et al.* discloses the reduction of the confidence score threshold level ((see page 10, right column, [0154], lines 17-19). It is inherent that the constant update and learning of the system presented in the reference would alter the confidence score threshold as it would alter the confidence level of the word.

7. Claim 7 is rejected under 35 U.S.C. 103(a) as being unpatentable over Kennewick *et al.* in view of Crepy *et al.* and Knott *et al.* as applied to claim 5 above, and further in view of Bickley *et al.* (US 7,013,276).

As to claims 7, Kennewick *et al.*, Crepy *et al.*, and Knott *et al.* disclose improving the performance of a speech recognition engine. However, Kennewick *et al.*, Crepy *et al.*, and Knott *et al.* do not specifically disclose the notification to a developer when the score is lower than a threshold value. Bickley *et al.* discloses a alert mechanism for words that are similar and are subject to confusion (see col. 10, lines 63-65) from threshold calculation (see col. 10, lines 38-40). It would have been obvious to one of ordinary skilled in the art to modify the speech recognition performance methods presented by Kennewick *et al.* as modified by the use of a notification sent to a software developer when value is below threshold. The motivation to combine these references involves the distinguishing between similar words, which may not be recognized by speech recognition engines (see Bickley *et al.* col. 2, line 27-36).

Conclusion

8. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Paras Shah whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-FRI. 7:30a.m.-5:00p.m. EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Xiao Wu can be reached on (571)272-7761. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2609

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

P.S.

05/11/2007


XIAO WU
SUPERVISORY PATENT EXAMINER